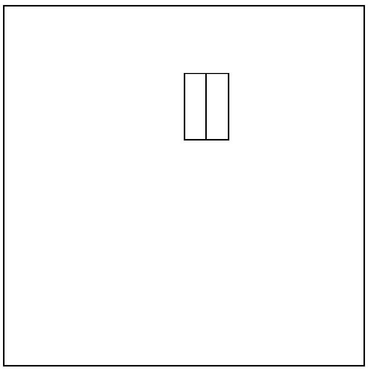
Project 2

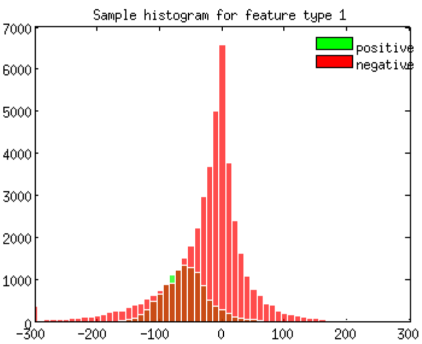
He Ma SID: 904434330

Part 1

For this project, I used the same 4 types of features as described in Viola and Jones's paper[1]. Below are the examples for each type of feature and the corresponding histogram for the distribution of positive and negative sampels.

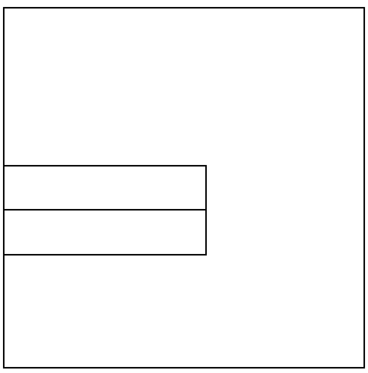
Type 1

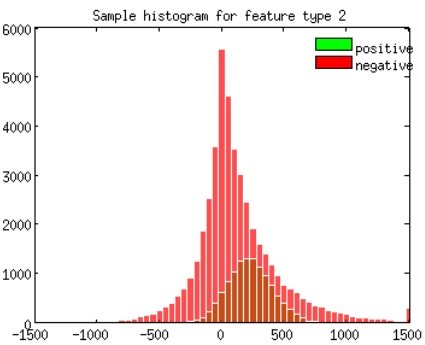




Feature = sum(right) - sum(left)
7680 features in total.
Threshold is around -29 for the example.

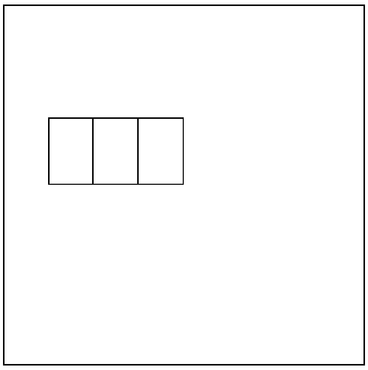
Type 2

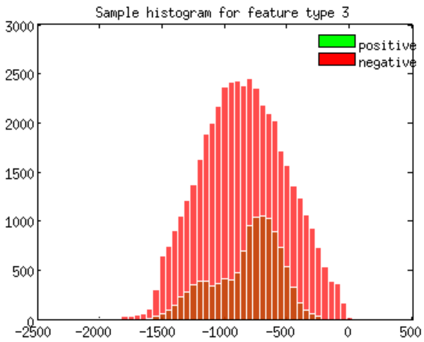




Feature = sum(top) - sum(bottom)
7680 features in total.
Threshold is around 100 for the example.

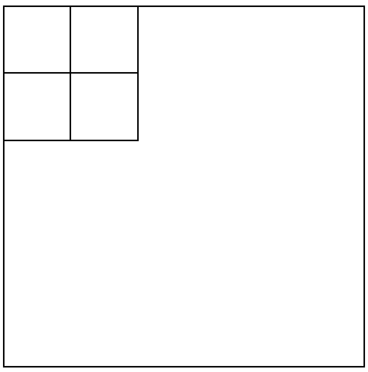
Type 3

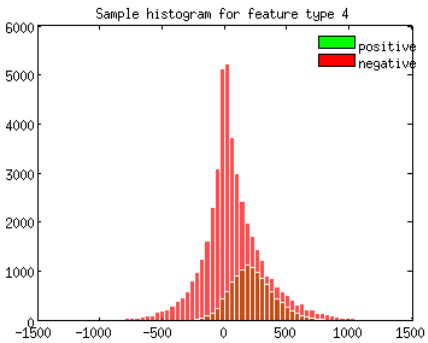




Feature = sum(center) - sum(left) - sum(right)
4800 features in total.
Threshold is around -700 for the example.

Type 4

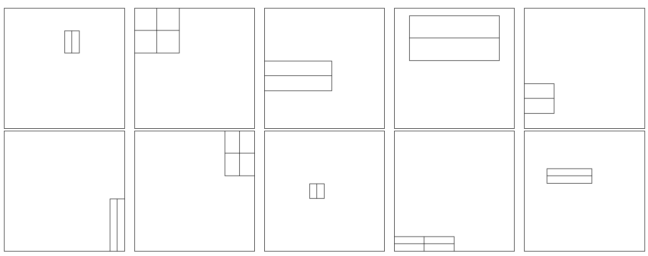




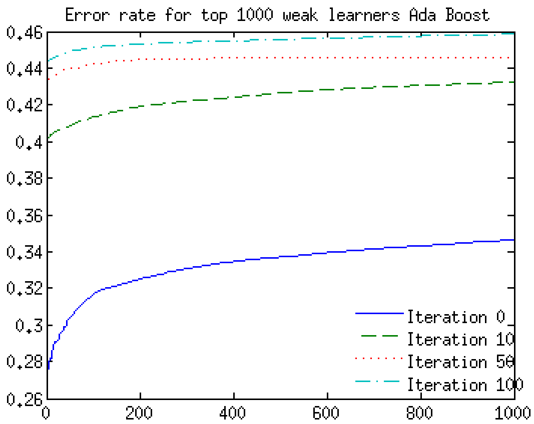
Feature = sum(top left) + sum(bottom right) - sum(top right) - sum(bottom left)
4096 features in total.
Threshold is around 100 for the example.

Part 2
I used the same algorithm as described in Viola and Jones's paper[1].
All 16 by 16 training images are used to train the model.
Training error rate: 4.9% for 100 iterations when threshold is 0

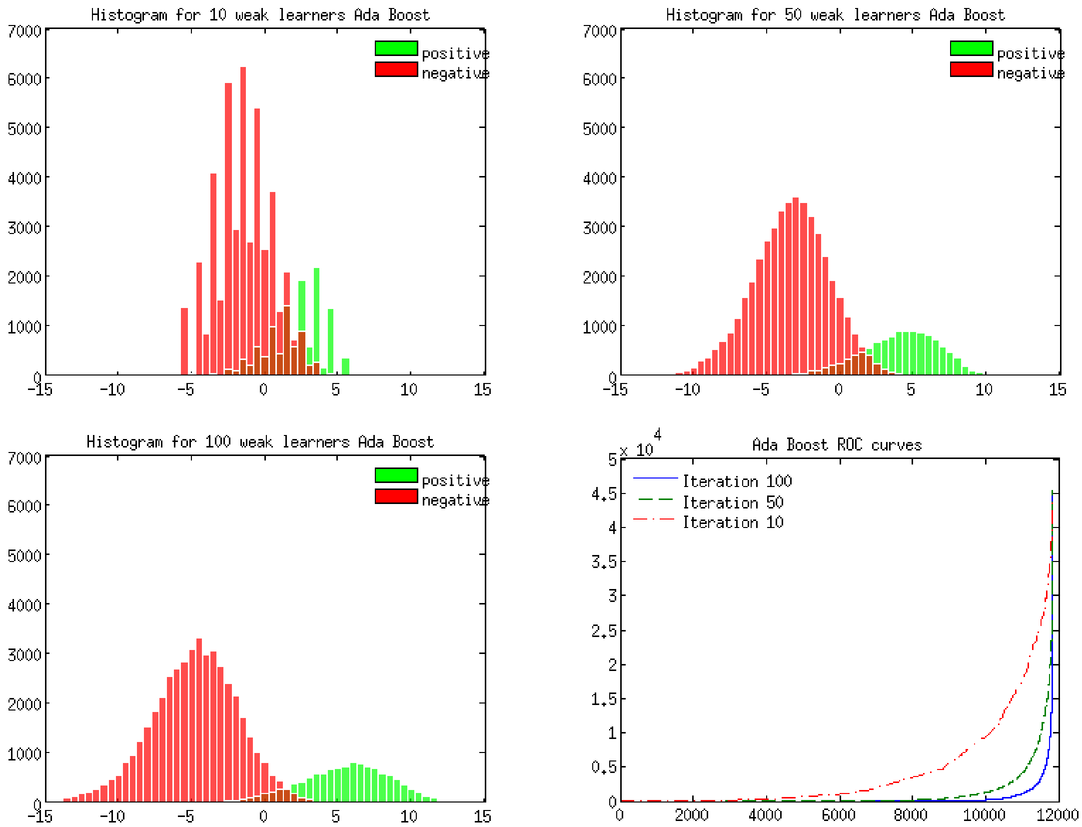
1. The top 10 features I got after 100 iterations of Adaboost are:



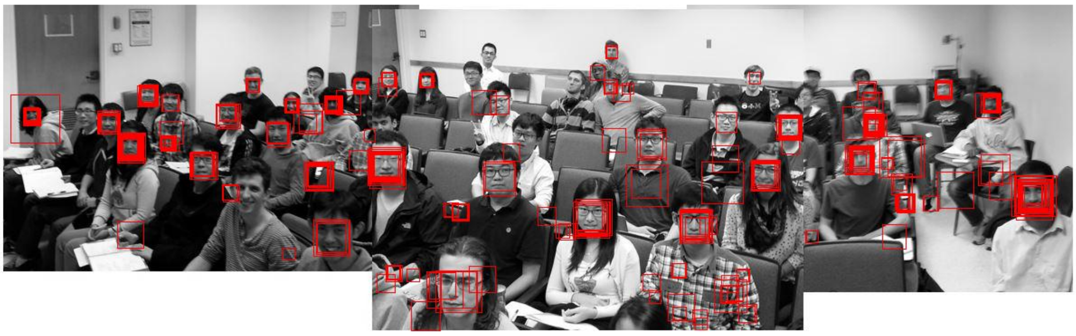
2. Blow are the curves for the errors to the 1000 weak classifiers at T=0,10,50,100. As we get more iterations, the error rates for the weak classifiers increases. More weak learners have error rate closer to 0.5 as we have more iterations.



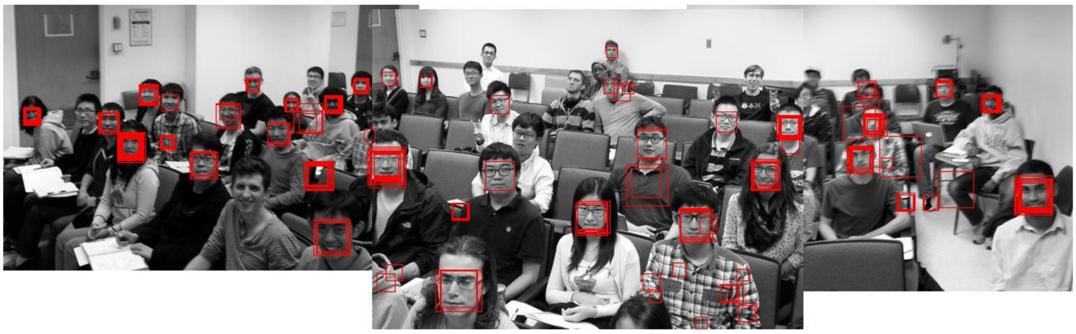
3. As we have more iterations, the positive and negative samples get more separated. As we can see from the ROC curve, the model is much better as we have more iterations.



4. Below is the test image using a threshold of 5.5 with scaling 0.3:0.05:1.3. The reason I'm choosing this threshold is that the strong learner returns a value less than 5 for all training non-face images. Because the number of non face fragments will be much bigger than the number of face fragments in a photos, we want to get as few false positive signals as possible.
   True Positive: 32
   False Negative: 11
   False Positive: Around 30
   True Negative: All other 400000 fragments



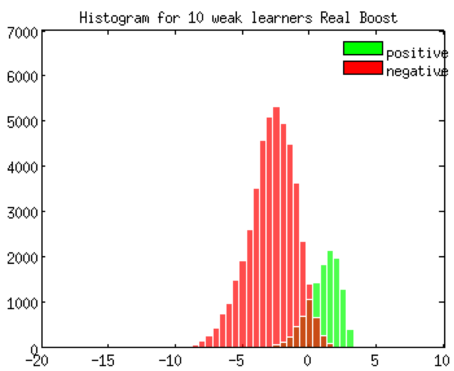
5. Adding false positive fragments of background image to the training set improves my Adaboost model slightly. There seems to be fewer false positives. But it is not significant. Most of the false positive of the class image do not come from classroom background, but people's clothes and other parts of body. The threshold chosen also influence the testing result. There is not a good way to determine whether the threshold choosen is optimal or not for both models. The picture shown below uses threshold 5 too. I believe the difference observed here comes mainly from the threshold chosen.

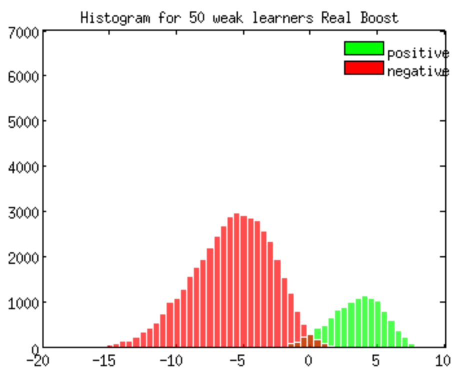


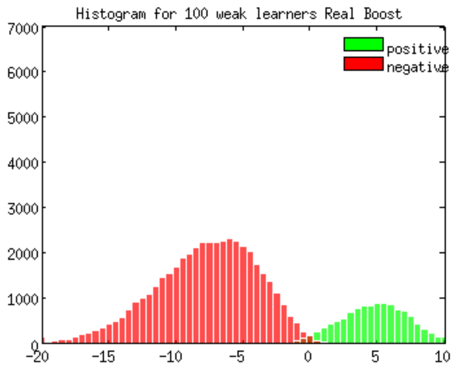
Part 3

I used the same algorithm as described in Wu and Ai's paper[2]
All 16 by 16 training images are used to train the model.
Training error rate: 0.94% for 100 iterations when threshold is 0

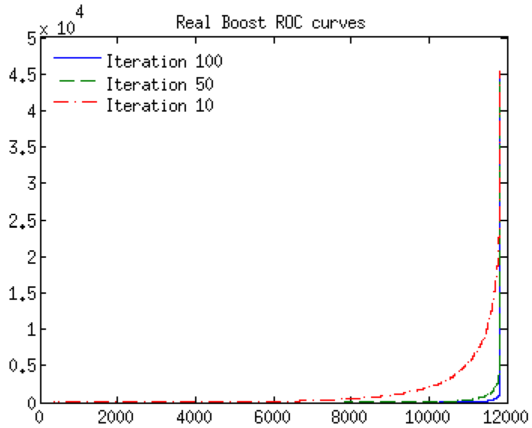
1. As we have more iterations, the positive and negative samples get more separated.



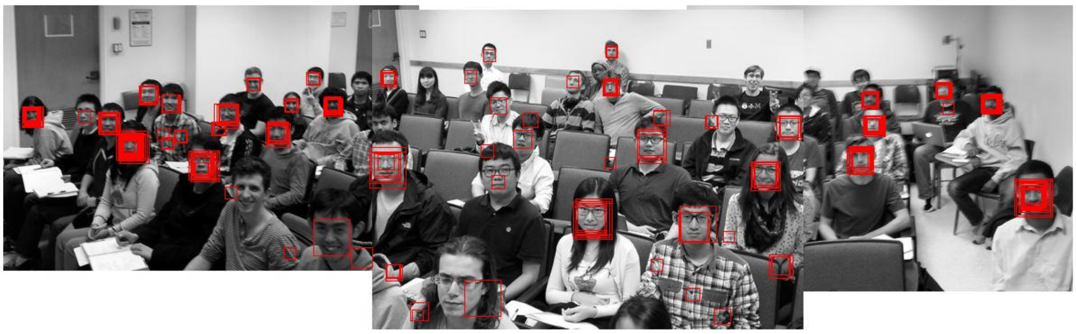




2. As we can see from the ROC curve, the model is much better as we have more iterations. The model seems to converge quickly. The model only improve slightly from iteration 50 to iteration 100.



3. Testing the model on the classe image using threshold as 3. The result is better in term of number of faces detected. There are more false negative as well. However, It think it depends more on the threshold choosen. By tuning the threshold, this model should achieve a better result.



Reference

P.Viola, M.Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, 2001
B.Wu, H.Ai, "Fast Rotation Invariant Muti-View Face Detection Based on Real Adaboost", 2004